Useful Relations for Countercurrent Distribution Computations

Expressions have been developed for the prediction and evaluation of results from countercurrent distribution experiments. Included is a method for determining the number of transfers required for a given "degree of separation" and pair of partition coefficients. An approximate relation is proposed for determining the number of transfers required to resolve composite peaks.

RECENTLY Craig (1, 2) has introduced several ingenious devices for conducting a series of liquid-liquid distributions between immiscible solvent pairs. The use of these new multiple distribution techniques, which Craig has termed "countercurrent distribution," is becoming increasingly popular as a means of separating, purifying, and identifying compounds. It has also been used as an analytical means of determining purity, inasmuch as many substances have essentially linear partition isotherms in a variety of solvent combinations and when this is true it is possible to estimate accurately the amount to be expected in a certain tube for a given partition coefficient.

The calculation of theoretical distribution curves has been given an excellent treatment by Williamson and Craig (7). Calculations of the distribution curves for experiments involving a large number of transfers by the methods suggested are still somewhat time-consuming, however, and it requires considerable familiarity with them to make even a rough approximation as to whether the method would be applicable to certain problems. It has been the author's experience that the possible utility of the method is grasped more readily when one has available an equation relating the known partition coefficients and desired degree of separation to the number of transfers required. With a relation of this type it is at the same time possible to gain an idea of the relative efficiency of the Craig method as compared with continuous multiple extraction processes exemplified by paper and partition chromatography and countercurrent extraction. The manner in which such a comparison could be made is contained in the work of Mayer and Tompkins (6) on the theory of exchange resin column separations. Their treatment of the distribution of substances on an ion exchange column can be shown to be essentially mathematically equivalent to Williamson and Craig's (7) treatment of distribution of substances in a Craig apparatus. Mayer and Tompkins also give a graphical method for the determination of degree of separation which could be applied to countercurrent distributions. The present treatment is particularly suitable for distributions within a Craig apparatus. However, its application to column separations is also possible within the limitations already discussed by Mayer and Tompkins.

In addition to deriving equations of the above-mentioned type, a number of other useful approximate relations are given to assist in rapid computations. Included are an accurate method for determining the areas under the tails of distribution curves by means of probability function tables, and a simple approximate method for determining how many transfers are required to resolve the peaks of distribution curves when the partition coefficients are known or can be estimated. The derived equations will apply to those distributions obtained by analyzing both phases which Craig (2) has termed the "fundamental" operation. A complete discussion of the various operations and types of distribution apparatus is given by Craig and Post (2).

EQUATIONS FOR INTERSECTION POINTS AND MAXIMA

Attention is first centered on obtaining expressions for intersection points and maxima of the distribution curves of a system of two components, a and b, where the fractional amounts of a and b in the mixture are expressed by F_a and F_b, respectively.

Consider an experiment in which the upper phase migrates while the lower phase remains stationary. The tubes are numbered 0, 1, 2, 3, . . ., r. The lower layers of each tube are filled, the sample is placed in tube 0, equilibration of the substance between layers is effected, and the top layer of tube 0 is transferred to tube 1. The same process is repeated n times, each time all top layers being transferred to the next highest tube in number. The binomial expansion, Equation 1, then gives the fractional amount of the mixture as represented by substance a, $T_{n,r}^a$ to be found in tube r after n transfers when the partition coefficient is K_a. A similar relation holds for substance b. The fractional amount of the mixture in tube r after n transfers would be the sum

$$T_{n,r}^a = \frac{n!}{(n-r)!\,r!}\left(\frac{1}{1+K_a}\right)^n K_a^r F_a \qquad (1)$$

of the fractions represented by substances a and b. The intersection point of the distribution curves for substances a and b is found by equating $T_{n,r}^a$ and $T_{n,r}^b$ and is shown in Equation 2. The quantity $\log F_a/F_b$ is readily seen to be negligibly small

$$r_i = n\left[\frac{\log\left(\frac{1+K_b}{1+K_a}\right)}{\log K_b/K_a}\right] - \frac{\log F_b/F_a}{\log K_b/K_a} \qquad (2)$$

except when F_a and F_b are considerably different. In deriving approximate relations for the intersection point, therefore, we shall neglect the effect of this term.

Williamson and Craig (7) proposed the use of the normal probability function as a means of approximating the amount of material in a given tube when the number of transfers is large, and in line with their use of this function Equation 3 gives an approximate expression for $T_{n,r}^a$. The quantity r_{ma} in Equation 3 is the tube at which the maximum amount of substance a is located after n transfers. The intersection point of the distribution curves for two substances, a and b, would then be given by Equation 4.

$$T_{n,r}^a = \frac{F_a}{\sqrt{2\pi n \frac{K_a}{(1+K_a)^2}}}\;e^{-\frac{(r-r_{ma})^2}{2n\frac{K_a}{(1+K_a)^2}}} \qquad (3)$$

$$r_i = \frac{nK_aK_b - \sqrt{n^2 K_a K_b - 4\beta K_a K_b + 4\beta}}{K_aK_b - 1} \qquad (4)$$

where

$$\beta = 2.303 \log \frac{F_b\sqrt{2\pi n K_a/(1+K_a)^2}}{F_a\sqrt{2\pi n K_b/(1+K_b)^2}}$$

It can be demonstrated that term β in Equation 4 is small in most cases and Equation 4 in which this term has been neglected, will serve as a good approximation for practical application—i.e., n is greater than 20, K_a are between 0.1 and 10, and F_a and F_b are not greatly different. It is interesting to note that the first term on the right-hand side of Equation 2 is approximately equal over

$$r_i = n\left(\frac{K_aK_b - \sqrt{K_aK_b}}{K_aK_b - 1}\right) \qquad (5)$$

the complete range of values of K_a and K_b to the right-hand side of Equation 5.

Another approximate relation for the intersection point between the curves of two substances a and b is shown in Equation 6. This equation is deduced by assuming the intersection point is halfway between the maxima of the two distribution curves where the position of the maxima are located by Equation 7, discussed in the following paragraph. This formula is particularly good when the partition coefficients (and F_a and F_b) are not greatly different.

$$r_i = \frac{n}{2}\left(\frac{K_a}{1+K_a} + \frac{K_b}{1+K_b}\right) \tag{6}$$

It will now be in line with our ultimate objective to obtain an expression for r_{ma}. Craig (1) proposed the approximate relation (Equation 7) from empirical considerations.

$$r_{ma} = n\left(\frac{K_a}{1+K_a}\right) \tag{7}$$

The derivation of this equation is given here in order to acquaint one better with the degree of approximation involved. It is as follows: Differentiate Equation 1 with respect to n after using Stirling's approximation. Set the derivative equal to zero and the expression (Equation 8) is obtained after simplification. When n is large with respect to r, the first and third terms cancel, which leads to Equation 7, where r equals r_{ma} in this case.

$$\frac{n+\frac{1}{2}}{n} + \ln\frac{n}{n-r} - \frac{n-r+\frac{1}{2}}{n-r} + \ln\frac{1}{K+1} = 0 \tag{8}$$

USE OF PROBABILITY FUNCTION TABLES

As previously stated, Equation 3 is the normal probability distribution curve of unit area in which the standard deviation σ is assigned the value $\sqrt{nK}/(1+K)$. The area (A_t^a) under the tail of the distribution curve for substance a (see shaded area in Figure 1) can be defined as
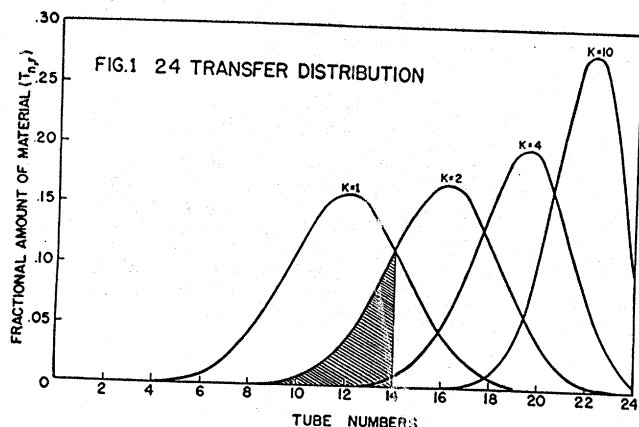
$$\int_{-\infty}^{-t_0} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt$$

FIG. 1 24 TRANSFER DISTRIBUTION

where t is equal to $(r - r_{ma})$ divided by σ_a. The limit of integration is designated as t_0 and variable of integration as t. This distinction, although mathematically precise, is not always made. However, it is particularly desirable in the present treatment.

Accordingly, multiplication of this area by the fraction, F_a, of a substance a gives the corresponding fractional amount of the total mixture as represented by substance a. The above integral is not commonly found in tables; therefore Equations 9 and 10 are given where the integral in Equation 9 can be found in Lange's handbook (4) and the integral in Equation 10 is tabulated in Hodgman's handbook (3). It will be useful to remember that for values of t equal to 1, 2, and 3, the corresponding areas under the tail are $A_1 = 0.159$, $A_2 = 0.0227$, and $A_3 = 0.0013$.

$$A_t = 1.0 - \int_{-\infty}^{t_0} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt \tag{9}$$

Table I. Comparison of Areas under Tails of Distribution Curves as Calculated by Equations 10 with 11 and 14 and Binomial Expansion

K	n	$r_i - r_m$	Area^a (Eq. 11)	Area^a (Eq. 14)	Area (B.E.)
1	24	6	0.012	0.007	0.0112
1	24	5	0.032	0.021	0.0318
1	24	4	0.076	0.051	0.0757
1	24	3	0.154	0.110	0.1536
1	24	2	0.271	0.207	0.2705
2	24	5	0.026	0.015	0.0191
2	24	4	0.064	0.042	0.0586
2	24	3	0.140	0.097	0.1375
2	24	2	0.258	0.193	0.2624
4	24	3	0.101	0.063	0.1145
10	24	2	0.144	0.078	0.1016
1	48	8	0.015	0.010	0.0139
1	48	4	0.155	0.124	0.1552

^{a} Areas calculated on basis of unit amounts of each substance.
^{b} Areas in last column are under those tails which point toward higher tube numbers. When partition coefficients differ greatly from 1, this specification is particularly necessary.

$$A_t = \frac{1}{2} - \int^{t_0} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt \tag{10}$$

For a reasonably accurate application of the equations for a continuous distribution to a discrete distribution as actually encountered, the value of $(r_i - r_{ma})$ used in calculating t_0 can be replaced by $(r_i - r_{ma} - 0.5)$. This arises from the nature of the discrete and continuous curves. This adjustment is particularly necessary when the number of transfers is comparatively small.

The use of Equations 10 and 11 in calculating areas under the tails of distribution curves is compared with results from the binomial expansion as shown in Table I.

$$\frac{r_i - r_{ma} - 0.5}{\sigma} = t_0 \tag{11}$$

ESTIMATION OF THE AMOUNT OF SEPARATION

With a means for estimating areas under the distribution curves, it is now possible to formulate a method for determining the degree of separation to be expected for two substances of partition coefficients K_a and K_b after a given number of transfers. As a measure of the amount of separation we use the area under the tail of the distribution curve of the substance of lowest partition coefficient from the point where the two curves intersect. The t value corresponding to this area will be called the degree of separation. Martin and Synge (5) have used t values in a similar sense in their theory of partition chromatography. Mayer and Tompkins (6) have defined and used t in an analogous manner in their theoretical analysis of exchange resin column separations.

The use of Equations 11, 5, and 7 leads to the following equations, where $K_b > K_a$.

$$nf - 0.5 = \sqrt{n}\left(\frac{\sqrt{K_a}}{1+K_a}\right)t_0 \tag{12}$$

where

$$f = \frac{K_a K_b - \sqrt{K_a K_b}}{K_a K_b - 1} - \frac{K_a}{K_a + 1}$$

$$t_0 = \left(\frac{1+K_a}{\sqrt{K_a}}\right)\left(\frac{nf - 0.5}{\sqrt{n}}\right) \tag{13}$$

When the factor 0.5 is omitted from Equation 11 we get Equation 14, which gives a rough estimation of the area which improves as n increases.

$$\frac{r_i - r_{ma}}{\sigma} = t_0 \tag{14}$$

Using Equations 12, 5, and 7, Equation 14 can be put in a more convenient form, as shown in Equation 15.

$$n = St_0^2$$

where
$$S = \left[\frac{K_a K_b - 1}{(K_b + 1)\sqrt{K_a} - (K_a + 1)\sqrt{K_b}}\right]^2$$

We now have two expressions for estimating the degree of separation for given partition coefficients and a given number of transfers. Equation 13 is necessary where one deals with a limited number of transfers and intermediate degrees of separation and Equation 15 is desirable for large numbers of transfers and high values of t_0.

Comparable agreement to that previously shown in Table I for calculated areas under tails of distribution curves can be expected using Equation 12, providing we round off the $(r_i - r_{ma})$ values to the nearest whole numbers, as must necessarily be the case in actual operation. Although Equation 15 gives a rough estimate as compared with Equation 12, its convenient form makes it useful as a quick means for determining the fractionating power of the method once two partition coefficients are known. Therefore, in Figure 2 the separation factor is plotted against the higher partition coefficient. If it is desired to know the number of transfers required for practically complete separation, multiplication by 3^2 gives the answer immediately. Different curves correspond to various combinations of K_a and K_b. In deriving the expression for S it must be remembered that we have adhered to the convention that K_b is greater than K_a; therefore the abscissa in the plot corresponds to different values of K_b compared with values of K_a which have been denoted by placing the latter as subscripts to S for a given curve. The curve approaches the value of K_a asymptotically. Results of calculations from Equations 13 and 15 are more exact the nearer the partition coefficients are to unity.

One point of interest arising in connection with the separation factor S is the following. It is tempting to measure the degree of separation by the ratio of the partition coefficients or a function of the ratio of the partition coefficients—e.g., Martin and Synge (5) used a procedure of this sort in their theory of partition chromatography. The ambiguity in such a procedure as applied to countercurrent distribution studies is clearly brought out in Table II, where the factor S is shown for different values of K_a and K_b corresponding to the ratio $K_a/K_b = 1.5$.

RESOLUTION OF PEAKS

Formulas are now available for answering another question concerning multiple distribution separations. Suppose we have two substances and their partition coefficients are known or can be estimated in some way. How many transfers would be required before the composite distribution curve would show two peaks?

It can be demonstrated that the appearance of two peaks is observable when the quantity δ as defined in Equation 16 is large enough to be experimentally detectable. We are assuming F_a and F_b to be equal, which gives a minimum value for our answer. The

$$\delta = (T^a_{n,r_{ma}} + T^b_{n,r_{mb}} - 2T^a_{n,r_i}) =$$

$$\frac{1}{\sqrt{2\pi}\sigma_a} + \frac{1}{\sqrt{2\pi}\sigma_b} e^{\frac{-(r_{ma}-r_{mb})^2}{2\sigma_b^2}} - \frac{2}{\sqrt{2\pi}\sigma_a} e^{\frac{-(r_i-r_{ma})^2}{2\sigma_a^2}} \quad (16)$$

second term on the right-hand side of Equation 16 can be rearranged with the help of Equation 17 to give
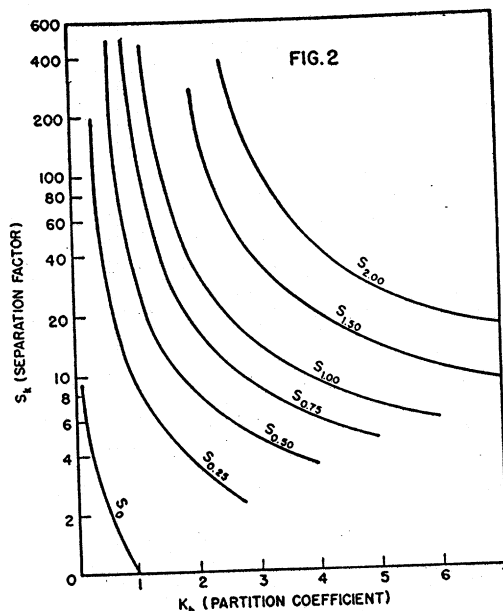
$$e^{\frac{-n(K_a - K_b)^2}{2(1 + K_a)^2 K_b}}$$

For even moderately different partition coefficients and large n this term is small enough to be neglected and the simple Equation 17 will suffice.

$$\delta = (T^a_{n,r_{ma}} - 2T_{n,r_i}) = \frac{1}{\sqrt{2\pi}\sigma_a}\left(1 - 2e^{\frac{-(r_i - r_{ma})^2}{2\sigma_a^2}}\right) \quad (17)$$

The quantity δ is approximately 5% of $T^a_{n,r_{ma}}$ when $r_i - r_{ma} = 1.22\sigma_a$, and this difference should be experimentally detectable.

$$d = 2(r_i - r_{ma}) = \frac{(K_b - K_a)n}{(1 + K_a)(1 + K_b)} \quad (18)$$

FIG. 2

It is now worth while to observe that $r_i - r_{ma}$ is one half the distance d (number of tubes) between the peaks as shown in Equation 18 where Formulas 5 and 7 were used. The above condition expressed in terms of n and the partition coefficients is shown in Equation 19.

$$n = \left[\frac{2.44\sqrt{K_a}(1 + K_b)}{K_b - K_a}\right]^2 \quad (19)$$

Therefore, Equation 19 gives us the number of transfers required to resolve a composite peak in terms of K_a and K_b.

Another helpful relation when dealing with composite peaks is shown in Equation 20. In this equation P_0 corresponds to the number of the tube at the peak or minimum between the peaks of a composite curve and P_e is where the peaks would be if $F_a = F_b$. Combined with $F_a + F_b = 1$, Equation 20 can be used in estimating the relative amounts of substances a and b where the

$$\frac{F_a}{F_b} = \left(\frac{K_a}{K_b}\right)^{P_0 - P_e} \quad (20)$$

partition coefficients can be estimated from the tails of the distribution curves.

Table II. Separation Factors for Partition Coefficients Corresponding to a Definite Ratio

K_a	K_b	K_a/K_b	S
0.6	0.4	1.5	110
0.9	0.6	1.5	99
1.5	1.0	1.5	98
6.0	4.0	1.5	172
9.0	6.0	1.5	230

LITERATURE CITED

(1) Craig, L. C., *J. Biol. Chem.*, 155, 519-34 (1944).
(2) Craig, L. C., and Post, O., ANAL. CHEM., 21, 500-4 (1949).
(3) Hodgman, C. D., "Handbook of Chemistry and Physics," 30th ed., pp. 204-8, Cleveland, Ohio, Chemical Rubber Publishing Co., 1946.
(4) Lange, N. A., "Handbook of Chemistry," 6th ed., pp. 257-60, Table XXVb, Sandusky, Ohio, Handbook Publishers, 1946.
(5) Martin, A. J. P., and Synge, R. L. M., *Biochem. J.*, 35, 1358-68 (1941).
(6) Mayer, S. W., and Tompkins, E. R., *J. Am. Chem. Soc.*, 69, 2866-74 (1947).
(7) Williamson, B., and Craig, L. C., *J. Biol. Chem.*, 168, 687-97 (1947).